Defence Research and   Recherche et développement
Development Canada   pour la défense Canada

DEFENCE **R&D** DÉFENSE

Computer-aided classification for a database of images of minelike objects

John A. Fawcett
Vincent L. Myers

Defence R&D Canada – Atlantic

Canadä

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **MAR 2005**	2. REPORT TYPE	3. DATES COVERED **-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Computer-aided Classification for a Database of Images of Minelike Objects	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Defence R&D Canada -Atlantic,PO Box 1012,Dartmouth, NS,CA,B2Y 3Z7**	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES **60**	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**			

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

This page intentionally left blank.

Computer-aided classification for a database of images of minelike objects

John A. Fawcett
and Vincent L. Myers

Defence Research & Development Canada Atlantic

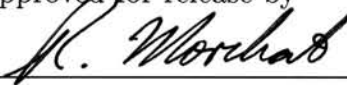Technical Memorandum

DRDC Atlantic TM2004–272

March 2005

Author

Joe Fawcett

Approved by

Ron Kuwahara

Ron Kuwahara
Head/Signatures

Approved for release by

R. Morchat

Kirk Foster
Chair/Document Review Panel

Abstract

In this report we describe some computer classification experiments with a database of sidescan sonar images. This database consists of 383 swaths of sidescan sonar data extracted by the authors from sea trial data collected over the last few years by DRDC Atlantic. The effects of the filtering and image segmentation processes on the resultant classification rates are considered. A number of kernel-based and nearest-neighbour classification schemes are examined. It is found that despite the complexities of the database considered in this report that high classification/low false alarm rates can be achieved.

Résumé

Dans le présent rapport, nous décrivons quelques essais de classification avec une base de données d'images sonar à balayage latéral. Cette base de données consiste en 383 fauchées de données de sonar à balayage latéral extraites par les auteurs à partir des données sur les essais en mer, recueillies pendant les quelques dernières années par RDDC Atlantique. Les effets des processus de filtrage et de segmentation d'images sur les taux de classification résultants sont pris en considération. On examine un certain nombre de systèmes de classification basés sur la méthode du noyau et sur le voisin le plus proche. On constate que, en dépit des complexités de la base de données examinée dans ce rapport, il est possible d'atteindre des taux élevés de classification et de faibles taux de fausse alarme

This page intentionally left blank.

Executive summary

INTRODUCTION

The use of computer-aided detection and classification techniques in mine-hunting is very important. The amount of data which is typically collected from a sidescan sonar during a minehunting survey is very large and thus it is hoped that computer algorithms can help a human operator with the workload. For autonomous vehicles the use of computer-aided detection and classification techniques may be even more important if (1) the vehicle is to change its survey on the basis of target detections (2) the amount of data to be transmitted, perhaps by underwater modem, is to be reduced.

RESULTS

It is shown that the values of classification features computed for sidescan images may depend significantly in some instances upon the prior filtering and segmentation algorithms used. Various algorithms will work better in different cases depending upon the target and seabed conditions. It is shown that a large number of features can be effectively used with a Kernel-based classifier. For the challenging set of sidescan images described in this report a classification rate of 90% could be obtained with approximately a 20% false alarm rate.

SIGNIFICANCE OF RESULTS

It has been shown that a totally automated approach of segmentation/feature extraction and classification can yield good classification and false alarm rates for sidescan sonar images from a database containing images with a large variety of objects and seabed backgrounds.

FUTURE WORK

It is hoped that a version of this classifier can be implemented for in-field trials. The sidescan sonar swaths used for this report were extracted manually. In the future we would like to combine this with a low-level automated detector providing the swaths for classification. Presently, computer-aided detection algorithms are being integrated with the Canadian Navy Route Survey Data Analysis Facility (RSDAF) software and it is hoped that the methods of this report will also be integrated with this software in the future.

Fawcett, J. & Myers, V., 2005. Computer-aided classification for a database of minelike objects. DRDC Atlantic TM 2004-272, Defence R&D Canada - Atlantic.

Sommaire

INTRODUCTION

L'utilisation de techniques de détection et de classification assistées par ordinateur pour la chasse aux mines revêt une grande importance. Étant donné la très grande quantité de données habituellement obtenues à l'aide d'un sonar à balayage latéral au cours d'une campagne de chasse aux mines, on espère que des algorithmes informatiques peuvent alléger la charge de travail de l'opérateur. Pour des véhicules autonomes, l'utilisation de techniques de détection et de classification assistées par ordinateur peuvent s'avérer encore plus importantes (1) si le véhicule doit modifier ses levés selon les détections de cibles (2) si la quantité de données à transmettre, peut-être par modem sous-marin, doit être réduite.

RÉSULTATS

Il est démontré que les valeurs des caractéristiques de classification calculées pour les images sonar à balayage latéral peuvent dépendre, de façon importante dans certains cas, des algorithmes de filtrage et de segmentation utilisés au préalable. Divers algorithmes seront plus appropriés dans différentes situations en fonction de l'état de la cible et du fond marin. Il appert qu'un grand nombre de caractéristiques peuvent être utilisées efficacement avec un classificateur basé sur la méthode du noyau. Pour le compliqué jeu d'images sonar à balayage latéral décrit dans le présent rapport, un taux de classification de 90 % pourrait être obtenu avec un taux de fausse alarme d'environ 20 %.

PORTÉE DES RÉSULTATS

Il a été démontré qu'une approche entièrement automatisée pour la segmentation et l'extraction des caractéristiques, ainsi que pour la classification, peut donner de bons taux de classification et de fausse alarme pour des images sonar à balayage latéral à partir d'une base de données renfermant des images qui comportent une grande variété d'objets et de fonds marins d'arrière-plan.

TRAVAUX À VENIR

On espère qu'une version de ce classificateur pourra être mise en uvre pour des essais pratiques. Les fauchées du sonar à balayage latéral utilisées pour le présent rapport ont été extraites manuellement. À l'avenir, nous aimerions combiner cette méthode avec un détecteur automatique de faible puissance fournissant les fauchées pour la classification. Actuellement, des algorithmes de détection assistée par ordinateur sont intégrés au logiciel du centre d'analyse des données de levés des fonds marins (Route Survey Data Analysis Facility,

RSDAF) de la Marine canadienne et on espère en outre que les méthodes
présentées dans le rapport y seront éventuellement intégrées.

Fawcett, J. & Myers, V., 2005. Computer-aided classification for a database
of minelike objects. DRDC Atlantic TM 2004-272, Defence R&D Canada -
Atlantic.

Table of contents

List of figures

List of tables

Acknowledgements

1 INTRODUCTION

There are a variety of stages in the processing and classification of sidescan sonar imagery. Initially the sonar data may be preprocessed - this includes the normalization of the data in the across-track direction to compensate for geometrical and attenuation effects, possibly scaling the data to lie within some range of values, and smoothing or median-filtering the data. The second stage in the process is often a low-level detection method. For example, this could consist of cross-correlating a set of simple highlight/shadow templates with the sidescan sonar data [1] and using a fairly low threshold to identify many possible minelike objects. This process will, in general, yield many false alarms, but it is hoped that many of these false alarms can subsequently be eliminated by a more sophisticated classification method.

To start the classification process, a small window around the detected object is defined. A number of features are then computed. Typically these features are defined in terms of the shadow and highlight regions of the image. For example, the length of an object's acoustic shadow on the seabed yields information about the height of that object. It is possible to define a number of features based on the dimensions and statistics of the shadow and highlight regions [2],[3],[4]. It is also possible to use the image pixels themselves, or the cross-correlation of the images with various templates as features [5].

In order to define the shadow/highlight/background regions of an image it is necessary to segment the image into these 3 regions. The subsequent accuracy of the feature extraction (for those features based on these regions) is highly dependent on this segmentation. Unfortunately, as will be seen, this is a non-trivial task and the optimal parameters for this segmentation may depend upon the seabed type and even the object itself.

In Ref.5 an approach was described which avoided the detailed segmentation of an image and instead used the cross-correlations of the images with an appropriate set of templates as the features. This method was very successful in the task of differentiating cylindrical targets (the MOG 5 cylinders, which are also in this database) from clutter. In this report, we will also consider this approach, but extended to the case of several target types.

Over the last few years, DRDC Atlantic has collected many sidescan sonar images of minelike objects placed upon the seabed. The sonar used was a Klein 5500 sidescan sonar. As part of the development of the DRDC Atlantic Sonar Image Processing System (SIPS) [6], tools were developed which allow an user to click on an object of interest during playback, and to extract a specified number (usually 21) of pings of data containing the target and associated

ancillary data. This tool was used by the authors to construct the database of swaths studied in this report. This process is a manual detection process. However, for the data set discussed below we made sure that various rocks, logs, etc were also included to test the classification process.

In June 2001, the Canadian Navy deployed 4 cylinders on the seabed of Herring Cove, Halifax and data was collected for these cylinders at a variety of aspects and across-track ranges. In July 2001, this site was revisited as part of the joint DRDC Atlantic/SACLANT Centre (now NURC) MAPLE trial [7] where in addition to the original cylinders, a number of additional targets, including dummy Mantas, additional cylinders, and moored spheres were deployed. In addition, many of these targets were also deployed at a site in St. Margaret's Bay. Also, as part of the trials of the Remote Minehunting System (RMS) off Esquimalt, B.C. a variety of well-known minelike objects were deployed. The database which is considered in this report used data from all these trials.

During the RMS trials, the speed of the towfish was approximately 8 knots. At this speed, typically all or 4 of the 5 beams of the Klein 5500 were used for each ping. The "despeckle" switch (an internal smoothing performed by the Klein beamforming) was off. For the Herring Cove data, the towspeed was usually about 4 knots which means that 2 or 3 of the Klein beams were "redundant": that is, they are not required to build up a complete image of the seabed as the towfish travels. Also, the "despeckle" switch was set at its lowest (non-zero) value. In general, because of the different geographical locations, the different deployment vessels and tow speeds, there is quite a variation in the seabed backgrounds and image quality for the swaths contained within this database. As well there is a large variety of the types of objects.

The swath files and database information which were formed from the data from these trials was distributed to the TTCP nations as part of the collaborative CAD/CAC project. Thus, we shall refer to this as the TTCP-database. We will use this database to examine some of the issues associated with feature extraction and the subsequent impact on classification. We will also investigate the use of template features which do not rely on a detailed segmentation of the object.

2 DESCRIPTION OF THE DATABASE

In the original definition of the swath image database considered in this report, the various objects were divided into different classes by the authors. These classes correspond to the different target types or types of clutter. The target classes were further subdivided according to the geographical location of their deployment. The different classes are described in Table I and the number of instances of swaths for each particular class is specified in the "Number" column. Some of the classes, such as the Mark 36 have only a few swaths in the database (8). This may be somewhat problematic during the training of the classifiers, as we do not insist in the random partitioning of the data for training and testing that there are a minimum number of occurrences for each class. Thus, it is possible that for some partitions of the training/testing set that there are no instances of a particular target type (especially one with only a small number of swaths) in the training set. The emphasis of this report is, in fact, on classifying the swaths as minelike or non minelike. Thus, all the target types will be assigned the label (+1) and all the clutter and curiosity swaths will be labelled as (-1). Due to their large image sizes, the Shipwreck class is not used in the present study. The labels ± 1 which are assigned to the swaths of the various classes for the purposes of this report are shown in Table I.

Class description	Dimensions	Number	New Label
Clutter: Various rocks and objects on the seafloor that are not mines	Various	112	-1
Horizontal Concrete Cylinder: Horizontal concrete cylinders	Unknown	82	1
Vertical Concrete Cylinder: Vertical concrete cylinders	Unknown	8	1
MK-56 RMSB3 T1: MK-56 cylindrical mines deployed during RMS trial- same as Class 7 below	2.75m by 59cm diameter	7	1
Manta RMSB3 T2: Manta shape deployed for RMS trial - same as Class 11 and Class 16.	1.02m base diameter by 0.45m height	10	1
MK-52 RMSB3 T3: A MK52 cylindrical mine.	1.72m by 40 cm diameter	9	1
MK-25 RMSB3 T4: MK-25 mine deployed for RMS trial - same as Class 10.	2.02m by 47 cm diameter	12	1
MK-56 RMSB3 T5: Same as class 3.	2.75 m by 59 cm diameter	11	1
MK-62 RMSB3 T6: MK-62 aircraft laid seabed mine.	1.65 m by 25 cm diameter	5	1
MK-36 RMSB3 T7: MK-36 cylindrical seabed mine.	1.71 m by 45 cm diameter	8	1
MK-25 RMSB3 T8: Same as Class 6	2.02 m by 47 cm diameter	7	1
Manta RMSB3 T9: Same as Class 4 and 16	1.02 base diameter x 0.45 m height	11	1
MK-62 RMSB3 T10: Same as Class 6.	1.65 m by 25 cm height	2	1
Shipwreck: Various shipwrecks.	Various	9	N/A
Curiosity: Various interesting objects	Various	12	-1
MOG5 Cylinder: Water-filled cylinder, Herring Cove	1.83 m by .61m diameter	51	1
Q260 Manta: Same as Class 2 and 9 but deployed at Herring Cove	1.02 base diameter X 0.45m height	19	1
Sphere: Spherical objects suspended in water column	0.45 and 0.62 m diameter, 1 to 4 m above seabed	8	1

Table 1: *A description of the swath images. The column $Number$ is the number of swaths in each class, column $NewLabel$ is the label assigned to each class for the study of this report.*

Figure 1: *Some photographs of some of the targets included in the database*

3 PREPROCESSING AND SEGMENTATION

As discussed previously we will assume that a small window about the initial detection is extracted. We use the visible beams: for example, if all beams are visible and there are 21 pings the extracted data would have an along-track length of 105. In the cross-track direction, 121 points before the detection and 171 after are used, for a total of 293 points. However, this number may be adjusted so that: (a) the first point corresponds to the seabed (not the water column) and (b) the last point is not outside the range of data points. There are images from moored spheres present in the database and although these are an interesting target of interest and are included in the study of this report, they pose some unique challenges to the automated processing used. They often have little highlight and a very long shadow (due to the rather high elevation of the target off the seabed) which is often separated by a large across-track distance from the highlight (if there is any). In this report, where we fix the size of our window about the target, part of the image of the sphere may extend past its limits.

The success of subsequent feature extraction and classification depend upon the output of the preprocessing and segmentation stages. There will be cases where the segmentation stage is poor and thus the computed features will be very inaccurate. Also, the optimal preprocessing and segmentation may depend upon the seabed type and the target itself.

One of the segmentation methods we use in this report is based upon the work of [8] which uses an iterative approach. A pixel is declared to be shadow(highlight) if its value is less(greater) than a specified threshold. However, this threshold is allowed to increase (decrease) depending on the local connectivity with other shadow(highlight) pixels. Algorithmically, the threshold is increased in steps and each pixel considered. However, after a sweep through the thresholds, the connectivity between the pixels has, in general, changed and one must sweep through the pixels again. This is continued for a maximum of 10 sweeps or until the segmented image does not change.

A similar approach is to define threshold levels and acceptable local connectivities (measured by using a convolution with a 3 x 3 filter of ones). The segmented image is then obtained by combining the images (one or zero) which have the required connectivities at the various thresholds. This is a non-iterative approach. For both these segmentation approaches there are various parameters to be defined: the starting threshold and the final maximum threshold (requiring the maximum connectivity) and for the second approach

we use 4 different thresholds with required connectivities of at least 2,3,4 and 6. One would expect the iterative method would allow for more "growing" of the shadow region. This is often the case, but not always. The non-iterative method does not require that the pixel value itself is low, just that the local connectivity is acceptable.

It is not clear apriori which segmentation method will yield the best results and what are the best values of the various parameters. In general, we expect that the most appropriate method will depend upon the characteristics of the seabed background and the target's highlight and shadow. Thus, the approach we take in this report is to compute a set of 38 features with respect to shadow and highlight regions and to repeat this basic set using 5 different segmentation methods. This yields 5 different sets of features. However, we emphasize that the definition of these features is the same for each of the 5 sets. It is only the segmentation method and the details of the filtering of the original data which vary and as a result the subsequent computed values of the features. In general, the segmentations used to determine the shadow regions range from segmentation methods (or parameters) defining shadows with only very low pixel values to methods which allow for rather loosely-defined shadow regions.

In this report we will always median filter our data (within the window as discussed above) in the cross-track direction. For Method 4 of segmentation, we will also normalize this data in the along-track direction by the median value of each along-track vector of values. This was done because we found it to be beneficial in the cases that the seabed surrounding the target was quite "bright". The threshold values for the first 4 segmentation methods are defined in terms of percentiles of the data values. For example, one might define the lower shadow threshold as the 10% values of the sorted data values. In this manner, the absolute level of the data under consideration should not be important. The fifth segmentation method (and the corresponding feature set, Set 5) used for the shadow segmentation is based upon determining a hard threshold which minimizes an entropy measure of the data (based upon an implementation used in the SIPS[6] system). After the image has been segmented, there will be, in general, many regions of shadow and highlight. A region labelling algorithm is then used to group the shadow (highlight) regions into connected, labelled groups. Our method allows for pixels a specified distance δ away to be defined as connected.

In the Appendix, we list the 38 basic features (non-templates) which are computed. As well, a summary of the filtering and segmentation parameters for the 5 filtering/segmentation methods is also given. As discussed above, the 5 sets of filtering and segmentation parameters which are used prior to the computation of the 38 features yields 5 sets of features which we will denote

as Set 1 - Set 5 respectively for the remainder of the report. We will refer to the 5 different filtering/segmentation methods as Methods 1-5. In some cases when we are discussing a particular feature value which results from the feature computation after the application of one of the Methods (e.g. Method 3), we will refer to the value as computed by, for example, Method 3.

4 TEMPLATES

As was discussed in [5] another set of useful features are the values of the cross-correlations between simple ray-generated templates of the targets and the image. A ray-tracing algorithm was written in Fortran to generate the model images of a target on the seabed. The algorithm uses a triangular facet model of an object. It determines the points on the seabed from which a line to the sonar intersects the object (and thus these points correspond to shadow). A simple cosine rule is used to compute the amplitude of the reflection from the object. For each target type these highlight/background/shadow images are computed for 3 towfish altitudes (12, 15 and 18 m) and for 13 possible seabed ranges $15 \leq r \leq 75 \ m$. For the cylindrical targets, the templates are also computed at 12 azimuths. For the 6 target types considered, a large number of templates are generated. These are all saved in a single file. During the feature computations, the templates for the nearest altitude and nearest range are extracted from the file. As well, the cross-correlations between the edge-masks, etc are also computed. In [5] only the MOG5 cylinders from Herring Cove were used. Here will use templates for (1) Manta class (2) vertical cylinder 0.8 m (high) x 0.6 m (diameter) (3) cylinder 2.75 m (long) x 0.6 m (diameter) (4) cylinder 1.8 m (long) x 0.6 m(diameter) (5) cylinder 1.65 m x 0.25 m (6) cylinder 1.7 m x 0.42 m. In Fig. 2 we show 3 of the facet models used in generating the ray highlight/shadow images.

For the template matching the image data is byte-scaled using the top $99th$ percentile of the data. The mean of this scaled image is subtracted off and only values below the 20% and above the 95% levels are used in the matching. In determining the optimal matching template, some care must be taken. First for each cylinder class, the cross-correlations between all the different aspects are computed. However, as was discussed in [5] it seems best to normalize the result by the L_2 norm of the template; this favours somewhat bigger templates over the smaller ones. (For smaller templates it is easier to get obtain a good cross-correlation value without necessarily overlapping most of the target image.) For the cylinder classes the optimal aspect template is then used for the computation of the additional features. For the Manta and vertical cylinder classes this aspect determination is not required. A total of 15 features for each target type is computed. These features include the value of the cross-correlation, the cross-correlation between the differenced template and the differenced data (in the across-track and along-track directions), the percentage of template shadow pixels overlapping data shadow pixels, etc. These features are outlined in the Appendix and the total of 90 template features are denoted as Set 6.
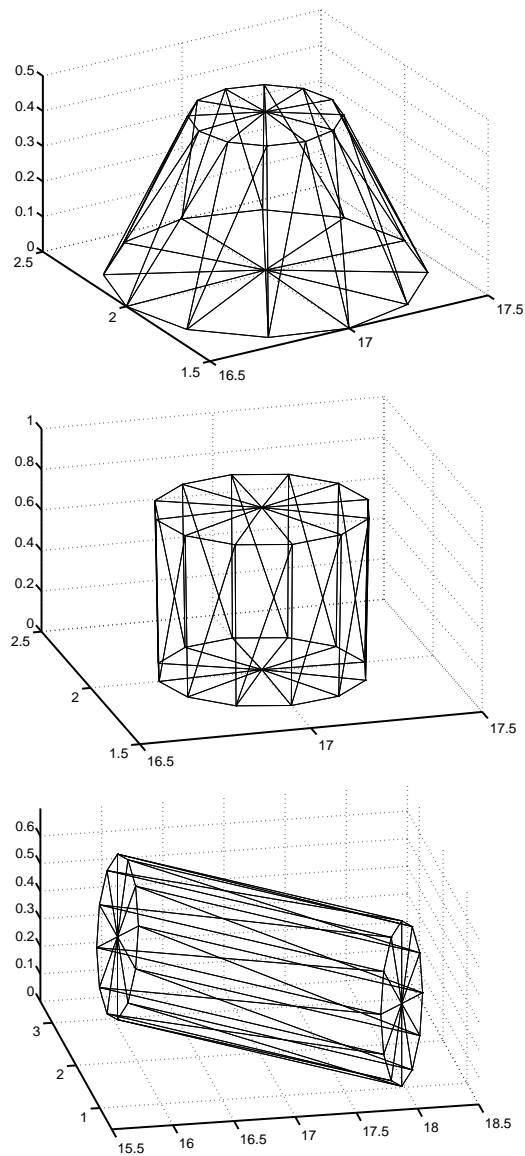
Figure 2: *Facet models for ray-tracing for (1) Manta (2) Vertical Cylinder (3) Mark 56 at angle of 45 degrees*

5   CLASSIFICATION ALGORITHMS

We have described a large number of possible features: there are 5 sets of 38 features (corresponding to using the 5 different sets of filtering/segmentation parameters, Methods 1-5) and an additional 90 template features. In the numerical examples, we will consider these as 6 sets of features. We will consider these sets of features individually and combined into larger sets. We will also consider combining the outputs from the classifier working on the individual sets. In this paper, we emphasize the basic two-class problem: target class or clutter. One can consider the multi-class problem as a sequence of 2-class problems, or after we have classed the object as a target, we could simply use the closest target class to classify it. The main classification methods we utilize in this report are Kernel-based classification methods [9,10]. We will also consider the simpler and computationally faster nearest-neighbour approach [11].

5.1   Kernel-based classifier

The basic concept of this type of classifier is quite straightforward. The set of features which constitute the training can be thought of as a set of vectors $\vec{f_i}$ where $\vec{f_i}$ denotes the set of features for swath i. There will also be a label -1 (clutter class) or $+1$ (target class) associated with it. The original features are then transformed into a new set of features by a non-linear mapping ϕ. We will not explicitly deal with these mapped features $\phi_j(\vec{f_i})$, as they are implicitly defined by the choice of the Kernel Function, which in general, will induce an infinite number of feature vectors (the eigenfunctions of the kernel operator) (note that j denotes the jth feature of the mapped feature vector $\vec{f_i}$). There are many possible Kernel Functions. The one we use is the exponential Kernel

$$K(\vec{x}, \vec{z}) = exp(-\|\vec{x} - \vec{z}\|_1/p), \tag{1}$$

where $\|\vec{x} - \vec{z}\|_1 \equiv \sum_{i=1}^{M} |\vec{x_i} - \vec{z_i}|$ and M is the number of features. We also tried the kernels

$$K(\vec{x}, \vec{z}) = exp(-\|\vec{x} - \vec{z}\|_2^2/p^2), \qquad K(\vec{x}, \vec{z}) = exp(-\|\vec{x} - \vec{z}\|_2/p) \tag{2}$$

but found that the kernel of Eq.(1) performed best for this particular problem.

A common classification approach in a two-class classification problem is to determine the plane in feature space which best separates the two classes. Below, we describe the method in the mapped domain $\vec{\phi}(\vec{x})$. Once the coefficients

of this plane (the vector $\vec{a}$) (and possibly a constant or bias term) have been determined to yield the discriminant function,

$$f(\vec{\phi}(\vec{x})) = \vec{a} \cdot \vec{\phi}(\vec{x}) + b \tag{3}$$

then the classification of a new swath (after the transformation to the new feature space) is based upon the sign of this discriminant function.

There are a variety of methods or criteria which can be used to determine the coefficients $\vec{a}, b$ in Eq.(3). Let us suppose that for the one class, we wish that $\vec{a} \cdot \vec{\phi} + b \geq 1$ and that for the other class $\vec{a} \cdot \vec{\phi} + b \leq -1$, then the distance between the 2 planes for which the equality, $\vec{a} \cdot \vec{\phi} + b = \pm 1$, holds is given by $1/\|\vec{a}\|_2$. The basic support vector classification problem is then

$$\begin{aligned} \min \ _{(\vec{a},b)} &\|\vec{a}\|_2 \\ \text{with} \ \ y_i(\vec{a} \cdot \vec{\phi}) &+ b \geq 1. \forall i \end{aligned} \tag{4}$$

This problem can be solved by using the method of Lagrange multipliers and Quadratic programming. In order to allow the optimizing solution of Eq.(4) to be more robust to noise one can introduce the concept of slack variables which allow the inequalities of Eq.(4) to be violated somewhat. After some derivation [9] it can be shown that the optimization problem in the dual space (i.e., after using Lagrange multipliers) is given by

$$\begin{aligned} \max \quad &\sum_{i=1}^{N} \alpha_i - 1/2 \sum_{i=1}^{N} \sum_{j=1}^{N} y_i y_j \alpha_i \alpha_j K(\vec{x}_i, \vec{x}_j) \\ \text{with} \ &\sum_{i=1}^{N} y_i \alpha_i = 0, \qquad 0 \leq \alpha_i \leq C \end{aligned} \tag{5}$$

where the number of swaths in the training set is N. In deriving Eq.(5) use has been made of the important relation that the inner product of the mapped features vectors,

$$\vec{\phi}(\vec{x}) \cdot \vec{\phi}(\vec{z}) = K(\vec{x}, \vec{z}). \tag{6}$$

Once the values of α_i have been determined the discriminating function is given by

$$\sum_{i=1}^{N} y_i \alpha_i K(\vec{x}_i, \vec{x}) + b \tag{7}$$

where the bias term b is found by the condition

$$\sum_{i=1}^{N} y_i \alpha_i K(\vec{x}_i, \vec{x}_j) + b = 1 \tag{8}$$

for values of j such that $0 < \alpha_j < C$. It is interesting to note that the discriminating function involves only those values of α which are non-zero (and the corresponding feature vectors are called the support vectors) . Thus only the data points which correspond to the these values of α have any influence on the future classification. Those training points which are more distant from the separating plane have no influence. If the number of "support" vectors is small then the required computations (and the storage for the data points corresponding to the support vectors) for classification is also small. This is certainly advantageous for very large size classification problems for which there may be thousands or even millions of datapoints used in the training.

In the above approach, the solution of Eq.(5) requires some form of a quadratic-programming algorithm. One can also use least-squares formulations to estimate the parameters α and b. In this case, we will simply minimize the L_2 error between the predicted label values and their true values, $\vec{Y}$. A regularisation parameter γ can also be included. There are 2 basic formulations, one including the bias term in the estimation and one not including it. In the first case, the system of equations has the form

$$\begin{pmatrix} 0 & 1 \cdots 1 \\ 1 & K + \frac{1}{\gamma}I \end{pmatrix} \begin{pmatrix} b \\ \alpha \end{pmatrix} = \begin{pmatrix} 0 \\ Y \end{pmatrix} \tag{9}$$

or where the bias term is not considered, simply

$$(K + \frac{1}{\gamma}I)\alpha = Y \tag{10}$$

Once α (and perhaps b) have been determined, the discriminating function has a form similar to Eq.(3),

$$f(\vec{x}) = \sum_{i=1}^{N} \alpha_i K(\vec{x}_i, \vec{x}) + b. \tag{11}$$

We have coded the 3 above algorithms, the first was coded in FORTRAN and used a general purpose optimization routine from the IMSL library[12] to solve the problem of Eq.(5) and the 2 least-squares approaches were coded in MATLAB. We found the performance was similar for all three methods and for the results shown here we use the least squares approach of Eq.(10).

There are three parameters which can be varied in the algorithm (besides the choice of the Kernel Function itself): (1) the scalar which is added to the diagonal of the Kernel matrix, (2) the value of p in the Kernel Function, (3) and a bias term, $\hat{b}$, when the discriminant function is applied. We found that the classification results did not depend significantly on the scalar added to the diagonal and we used a value of 0.005 in all our computations. The performance of the classifier can vary significantly with the parameter p and we will optimize the choice of this parameter by minimizing the classification error as averaged over random partitions of the data set into training and testing sets. The last parameter is not obvious. However, by varying the constant $\hat{b}$ from negative values (e.g. -1.5) to positive values (e.g 1.5) in the classification test (recall that in the determination of the weights α_j that $\hat{b}$ is assumed equal to zero)

$$\sum_{i=1}^{N} \alpha_j K(\vec{x}_i, \vec{x}) \geq \hat{b} \tag{12}$$

one can compute a ROC curve (probability of detection vs. probability of false alarm).

5.2   Nearest-neighbour classifier

This approach is a standard, simple classifier [11]. One considers a set of known swaths (training set) with their computed features. A new swath is then assigned the label of the closest swath with respect to the distance (using some distance measure) between the vectors of features. It is not difficult to see that this estimator can be considered as the limit as $p \to 0$ of the exponential Kernel-based classifier. As $p \to 0$ the matrix used in the training set becomes diagonal and the determined weights simply become the labels of the members of the training set. For a new swath, the computed exponential weightings to the elements of the training set will become increasingly dominated, as $p \to 0$, by the exponential with the smallest distance and thus this new swath will obtain the label of the closest swath in the training set (multiplied by the exponential weighting term). A generalization of the nearest-neighbour approach is to consider the N-nearest swaths in the training set and take the label of the majority. We will use N=1 and N=3 in our simulations.

6 NUMERICAL RESULTS

6.1 Some examples of segmentation and feature computation

In Fig. 3 we show 4 example images from the database, the resulting segmented images and the corresponding optimal template. This is the template from all possible templates which has the maximum cross-correlation with the data when normalized by the L_2 norm of the template. The segmentation results shown are for the fourth set of shadow/highlight segmentation parameters (Method 4 as described in the Appendix). The highlights shown were determined using a method whereby any highlight clusters whose centres fell within an acceptable distance of the leading edge of the shadow are accepted as highlight. For these swaths, it can be seen that the segmentations and the optimizing templates are good. There are other swaths for which the segmentations or optimizing templates are not as good (it should be noted that this optimal template is not used in the classification code, the features consist of the cross-correlations with all the target types). For example, for some of the Manta images an endon cylinder provides a good template match with the image.

In Fig. 4 we show the segmentations of the Manta of Swath 8 using the 5 different segmentation methods. The various segmentation methods use different parameters in terms of the length of median filtering (in the across-track direction), the thresholds for the shadow and highlight segmentation, and whether the iterative or non-iterative shadow determination was used. The fifth segmentation method used an entropy criteria for determining a hard threshold to use with the segmentation. Basically Method 1 considers only pixel values less than the 6% percentile as potential shadow pixels, whereas Method 4 has the loosest constraints on the shadow pixels. The details of these methods are given in the Appendix. There is a small highlight in the shadow region of the Manta. The shadow segmentations of Method 1 and 3 do not "make it" past this highlight whereas the other 3 methods do. Based upon the classification results discussed later, we found that Method 4 provided the best feature values. However, there are certainly examples where one of the other methods provides a better segmentation. In Fig. 5 we show the segmentation results for a dummy Manta deployed in Herring Cove. In this case the background contains a small linear shadow which is connected to the Manta shadow. This causes all the methods with the exception of Method 1 to associate this seabed
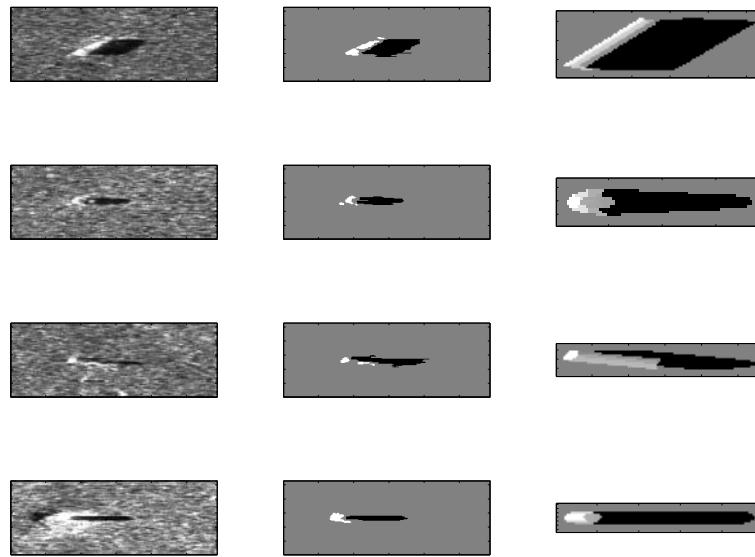
Figure 3: *Some representative target images and their segmentations using segmentation Method 4 and the automatically determined templates*

feature with the target.

The reason for using different parameter settings is to handle the effect of varying seabed and target conditions. In some instances the shadow/background contrast is not good and one needs to allow for a higher threshold on the shadow values to allow the shadow region to be sufficiently large. Also, the length of the median filter applied seems to have a significant effect. By increasing this length, a greater amount of speckle of the shadow regions can be reduced, thus improving the performance of the segmentation. It is hoped that by employing a variety of filtering and segmentation schemes, that at least one of the methods will produce good feature estimates. Of course, it is not known apriori which method will yield the best feature estimate and it is also not clear how to best utilize these various values for a particular features.

We now look at some of the feature values for some of the targets in detail. First we consider the MOG5 cylinders of which there are 50 instances (according to Table 1 there are 51 but one swath was rejected because it was at a different resolution setting). In Fig. 6 we show the computed lengths and heights of these cylinders after the application of the 5 filtering/segmentation methods. As can be seen the first method tends to underestimate the height somewhat, while the values corresponding to Method 5 are poor in some cases. Methods 2, 3, and 4 all provide reasonable estimates.

The estimation of the object length is somewhat problematic in the case of
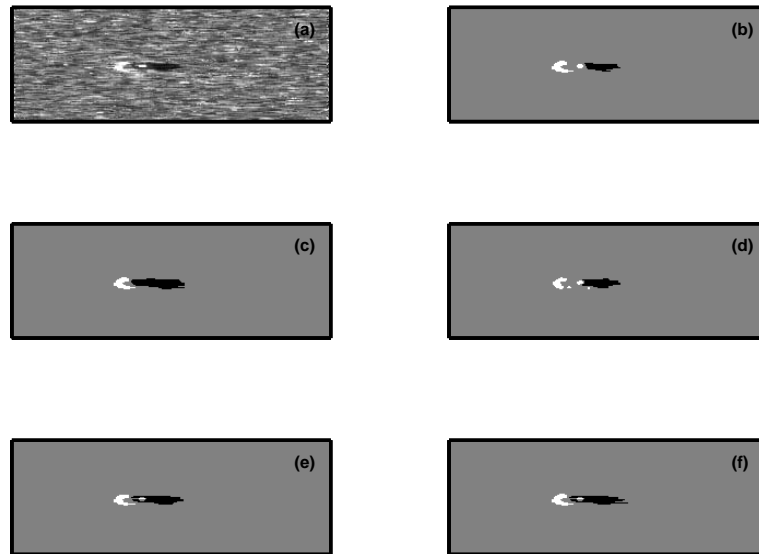
Figure 4: *The 5 different segmentations for the Manta of Swath 8, (b)-(f), original filtered image shown in (a)*
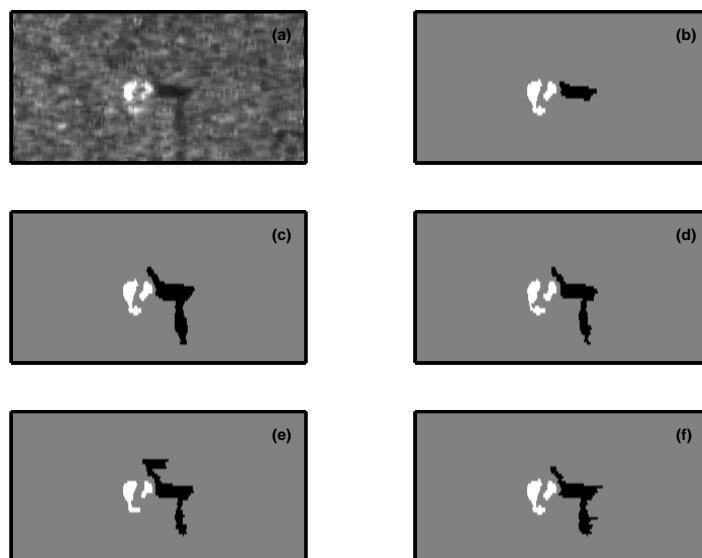
Figure 5: *The 5 different segmentations for the Manta (Herring Cove) of Swath 314, (b)-(f), original filtered image shown in (a)*
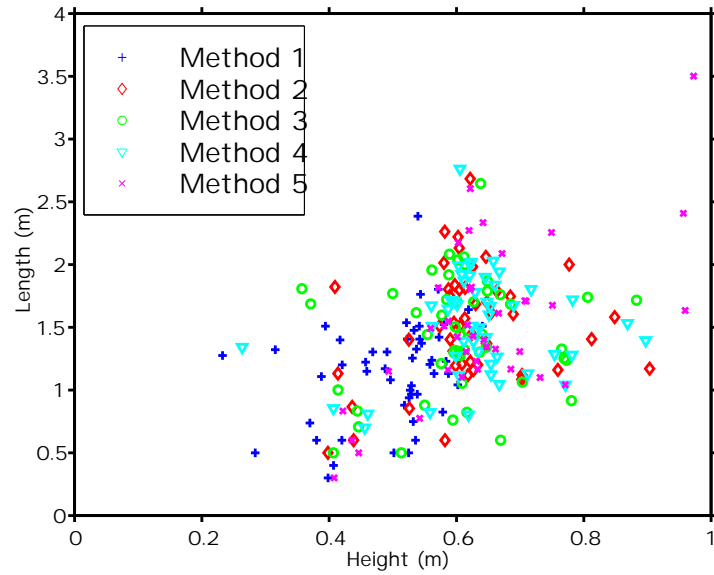
Figure 6: *The estimated heights and lengths for MOG 5 cylinders (0.61 m high x 1.83 m long) using the 5 sets of segmentation parameters (Methods 1-5 described in the Appendix)*
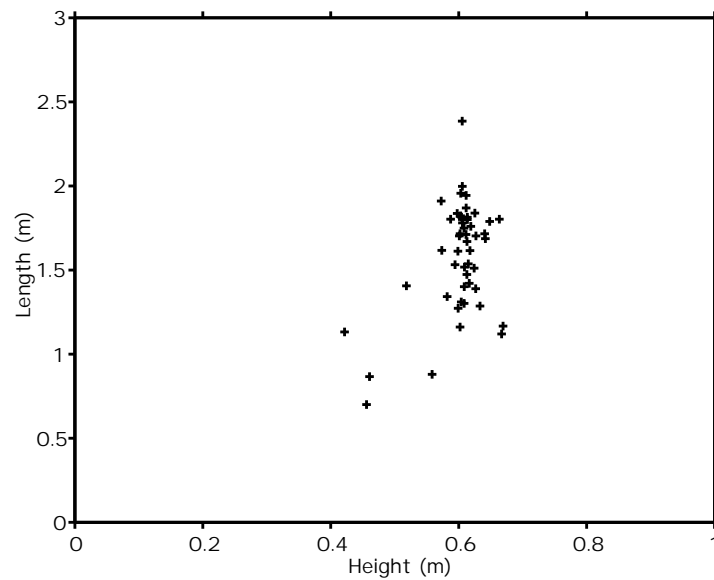
Figure 7: *The estimated heights and lengths for MOG 5 cylinders (0.61 m high x 1.83 m long) using the optimal of the 5 features for each swath*

cylindrical objects. For cylinders with a significant aspect it is necessary to try to account for the difference in the across-track pixels (for example, if we consider the leading edge of the shadow). This estimate can be difficult however if, for example the leading edge is "ragged". One can also use the highlight to estimate the length. However, there may also be problems in obtaining a reliable highlight segmentation. For the plots shown here we are using the length as computed from the leading edge of the shadow. Using Method 4, the computed mean height of the MOG5 cylinder is 0.641 meters (standard deviation of 0.118 m) and a mean length of 1.51 m with a standard deviation of 0.399 m. The height estimate agrees well with true height (diameter) of 0.61. The mean length is a little small and the standard deviation relatively higher than for the height. In Fig. 7 we show the distribution of heights and lengths obtained by selecting from the 5 methods, the estimates which are closest to the true values. This does significantly reduce the distribution of the heights (with respect to the values of, for example, Method 4), but does not significantly reduce the scatter of the lengths. The results of Fig. 7 are unrealistic in the sense that a computer algorithm does not "know" which is the best segmentation method to use for a particular swath. However, the figure does indicate that part of the scatter in the results of Fig. 6 may be due to not using the best set of filtering/segmentation parameters prior to computing the heights and lengths.

In Figs. 8 and 9 we repeat these computations for the Manta targets, two of which were deployed in the Esquimalt trials and all the instances of Manta targets from Herring Cove. Here there is a relatively large distribution in the height estimates. One reason for this is that it seems that the second Manta in Esquimalt may not have been lying flat on the bottom (perhaps, the deployment cable was pulling upwards on it) and the heights from this location seemed consistently too large. The mean height using Method 4 was 0.589 metres with a standard deviation of 0.147 m and a mean length of 0.872 meters with a standard deviation of 0.273 m.

In Figs. 10 and 11 we show the same results for the Mk 56 target. This was the largest minelike target and the estimates from Method 4, mean height .679 m (standard deviation .101m) and mean length 2.44m (standard deviation 0.383 m), are good (compared to the true values of 0.6m and 2.75m, respectively). Finally in Fig. 12, we show the height and length distribution of all the objects in the database which were labeled as clutter or curiosity. It can be seen that this class has a very wide distribution in height and length and thus it is not possible to eliminate many of the clutter/curiosity swaths on the basis of simple dimensions alone.
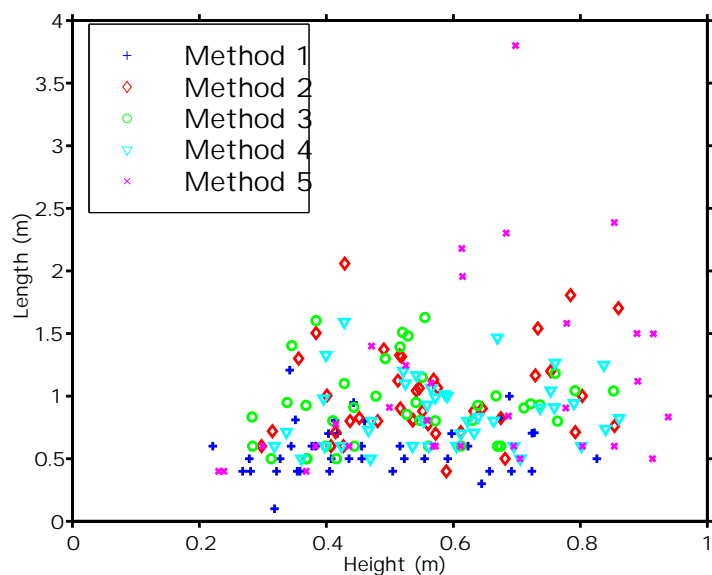
Figure 8: *The estimated heights and lengths for Manta shapes (0.45 m high x 1.02 m diameter) using the 5 sets of segmentation parameters*
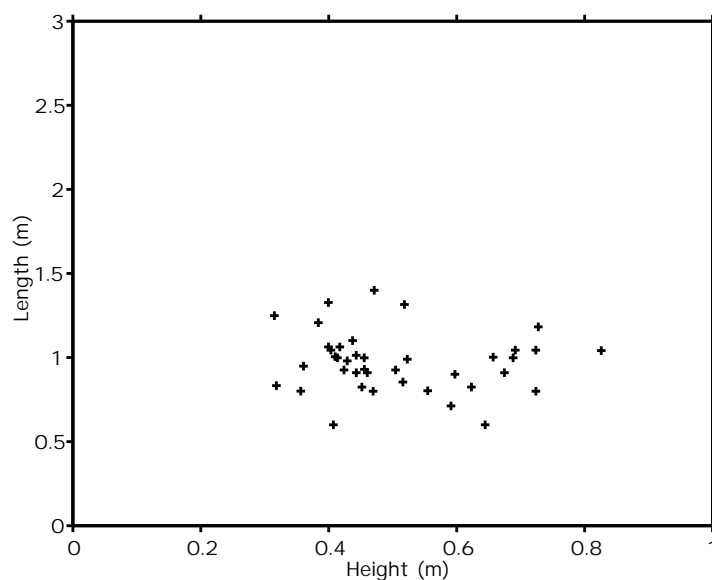
Figure 9: *The estimated heights and lengths for Manta shapes (0.45 m high x 1.02 m diameter) using the optimal of the 5 estimates for each swath*
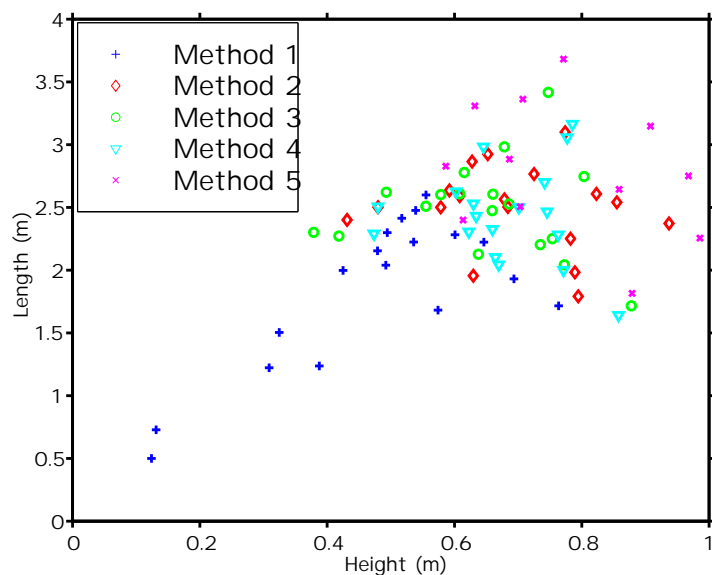
Figure 10: *The estimated heights and lengths for Mark 56 targets (diameter 59 cm x 2.75 m length) using the 5 sets of segmentation parameters*
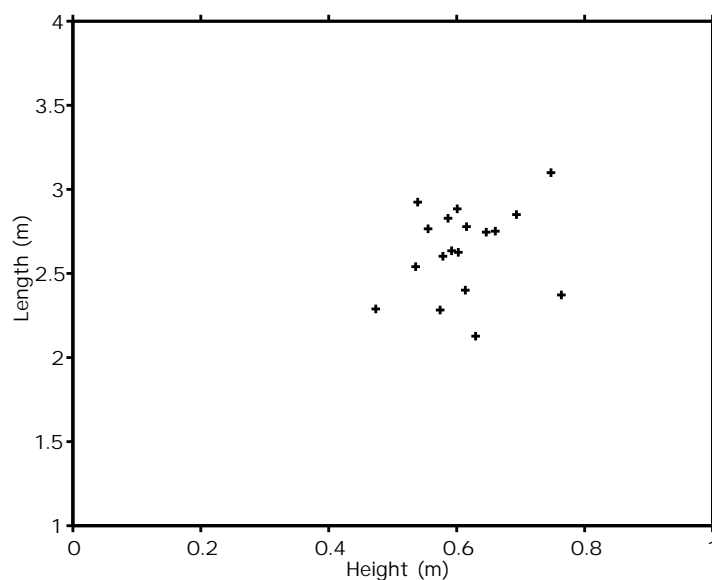
Figure 11: *The estimated heights and lengths for Mark 56 targets (diameter 59 cm x 2.75 m length) using the optimal of the 5 estimates for each swath*
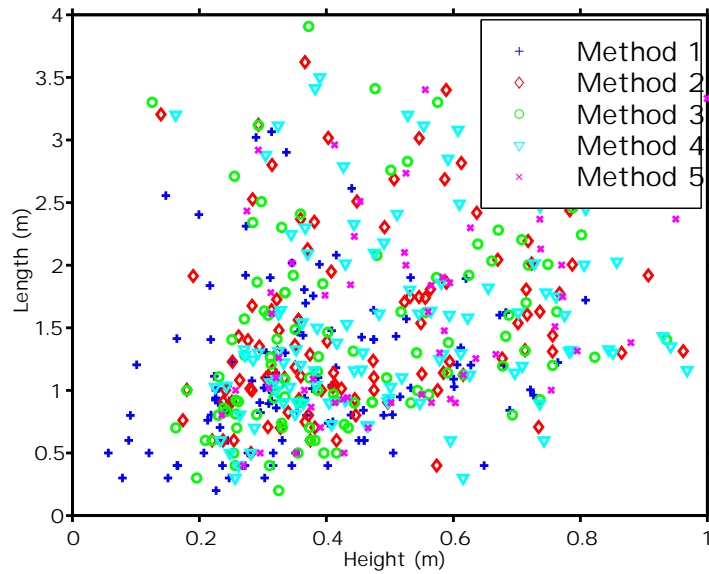
Figure 12: *The estimated heights and lengths for clutter (and curiosity) objects using the 5 sets of segmentation parameters*

6.2   Classification Results

We now consider, in more detail, the classification of swaths using various sets of features. As discussed in section 5 we will use Kernel-based feature sets with a least-squares method. As well, we will also give the error rates for the $N = 1$ and $N = 3$ nearest neighbour classifiers. We will consider a variety of different combinations of features and also the combinations of the outputs from classifiers. First, we consider the 6 sets of features separately: the first five sets are the sets of 38 features obtained after using the five different segmentation schemes and the set of template-based features. We also consider a set constructed from combining: (a) the first 4 sets of features (Set 7), (b) a set from the first 4 sets of features and the template features (Set 8), (c) Set 4 and the template features (Set 9), (d) all the features (Set 10), and finally (e) 2 sets (Sets 11-12) of features which are determined using Backwards Features selection [11].

In order to compute the following results, we first read in 383 swaths and their feature files. In the processing, some swaths which were large in size were skipped and we do not include any shipwreck swaths in this study. Any swaths which were identified as "Curiosity" were considered as a clutter event as well as those originally identified as " Clutter". All other swaths were considered as "Target". This includes the spheres and the horizontal and vertical concrete cylinders. Thus there are a wide variety of target types. After the elimination of the shipwreck files and the other swaths which were originally too large,

there was a total of 371 swaths to consider. This feature set and associated label were then randomly partitioned into 246 files for training and 125 files for testing. The features for this training set were then demeaned and normalized to have unit length. The computed feature means and normalizations were then applied to the testing set. After training the classifiers, the testing stage provided the number of targets which were misclassified as clutter (missed targets) and clutter misclassified as target (false alarm). As well, there is the total error rate. The missed target and false alarm rates were computed by summing all the missed targets and false alarms over all the simulations and normalizing these numbers by the total number of target (label 1) and clutter (label -1) events over all the simulations. There are, in fact, more targets than clutter in this dataset so that the total error rate is dominated by the missed target rate. The random partitioning of the files is repeated many (e.g. 301) times to yield fairly robust estimates of the error rates. We used the same partitionings for the different specified sets of features to compute the performance of several classifiers using the same set of simulations. We made no attempt during the partitioning to make sure that certain percentage of targets was represented. Thus, for some of the target classes with only a few instances, there could be particular partitions where none of them were in the training set.

We used backward feature selection to select an "optimal" set of features. In order to do this we start with the set of the first 280 features and start dropping the features in sets of 4 consecutive features (i.e, there are originally 70 groups of 4 features). This is clearly suboptimal in terms of finding an optimal feature set, but we did this to speed up the computations. We average the error results over 21 random partitions of the training and testing sets. In the discrimination test we set a threshold of 0.05 in an attempt to lessen the false alarm rate of the classifier at each stage. The parameter p was linearly decreased as the number of features decreased using the formula p = No. of features/280 ×1.5. The set of N-4 features which has the lowest average error is kept and then the process of determining the next set of 4 features which can be discarded is continued. In Fig.13 the average error rate as a function of the number of features discarded is shown. As well an optimized set was determined for the $N = 1$ nearest-neighbour classifier and that curve is also shown. From Fig. 13, it can be seen that the Kernel-based classifier has a minimum after about 200 features have been deleted, resulting in a set of 80 features. It is interesting to note that in this set there is a distribution of features from the different sets of features; there are 2 from Set 1, 10 from Set 2, 8 from Set 3, 20 from set 4, 16 from Set 5 and 24 from the template features. The nearest-neighbour classifier curve has a minimum at about 160 features deleted, resulting in a set of 120 features. Once again, this optimized set has
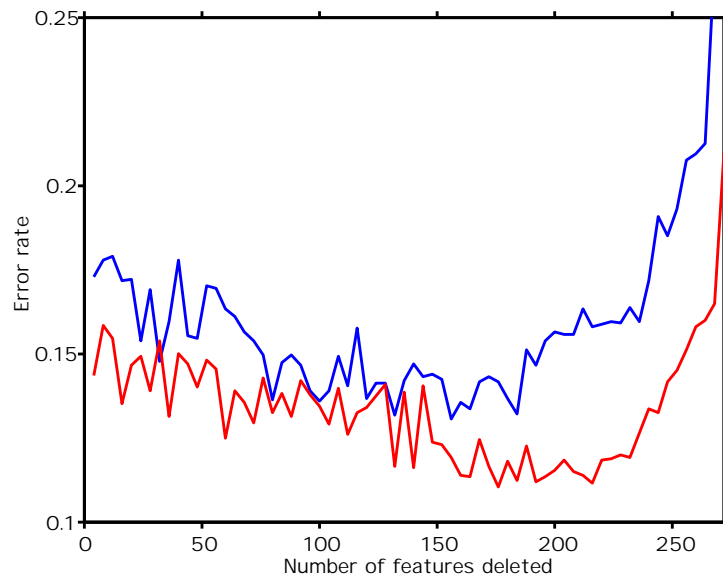
Figure 13: *The error rate as a function of the optimizing feature set size for Kernel-based approach, red, with a decreasing value of p starting at $p = 1.5$ and blue, the N1-nearest neighbour classifier*

a distribution of features from Sets 1-6. It can be seen that the kernel-based curve has a lower minimum error rate than the nearest-neighbour classifier. However, as will be seen the performance of the nearest-neighbour classifier, using its optimized feature set, is very good relative to classifiers using other features sets.

To investigate the performance of the kernel-based classifier as a function of the parameter p in the exponential kernel, we simply varied p and used 81 Monte Carlo simulations with respect to the partitioning of the training and testing sets to compute an average error rate for each feature set as a function of p. The error rate curves for Sets 4,6,8, and 11 are shown in Fig. 12. As can be seen, there are in fact 2 local minima in the curves. The optimized feature set, Set 11, has a significantly lower error rate than the other feature sets. Set 8 is the next best with a slightly smaller error rate than Set 4. The feature set 4 which results from using the filtering/segmentation parameters of Method 4 is the best of the individual feature sets (Sets 1-5, and the template features set 6). The feature sets 7-10 are all large feature sets with, for example, set 10 having 280 features and it is interesting to note that this not caused a deterioration in the classification results as compared to the smaller set, set 1-6. However, it is clear that the optimized feature set, set 11, definitely yields the best classification results. It is interesting to note that in the limit as $p \rightarrow 0$ the Kernel-based classifier should approach the nearest-neighbour
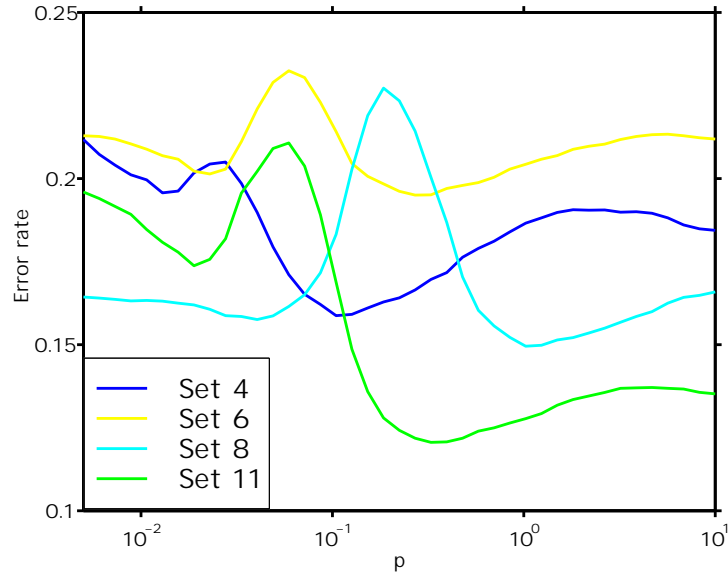
Figure 14: *The error rate for 4 sets of features as a function of p*

classifier. That is, a new swath is simply assigned the label of the swath in the training set which is closest with respect to the distance measure used in the exponential. Thus, for example, for Set 8 where the error rate for small values of p is quite close to the optimal error rate, we expect a nearest neighbour approach to work well.

Using the optimal values of p for each Feature Set we can compute ROC curves by varying $\hat{b}$ in Eq.(12) between -1.5 and 1.5 with 801 steps. For example, for $\hat{b} = -1.5$ then almost all swaths are classed as targets. The resulting ROC curves for features sets 4,6,8, and 11 are shown in Fig.15. From the data for these curves, the probability of false alarm for different probabilities of detection can be determined. In Table 2 below we give the false alarm rates for the feature sets for probabilities of target classification of 80%,90% and 95%. The results for feature Set 12 (the optimized set for the nearest neighbour classifier) are not shown. The optimal value of p for this set was 0.013 and because of this small value, many of the discriminant values (i.e. predicted label values) for the test set were very close to zero and the resulting ROC curve appears discontinuous for the spacing of p we used in the computation of the curves. It can be seen from this Table, that the classification rates using feature set 11 are very good, with a false alarm rate of 26.2% for a 95% target classification rate. Feature sets 8 and 10 were the next best sets. Of the individual feature sets, Set 4 was particularly good.

From the data used in the computation of the ROC curves we can find the smallest overall total error rate for each of the classifiers. These rates are
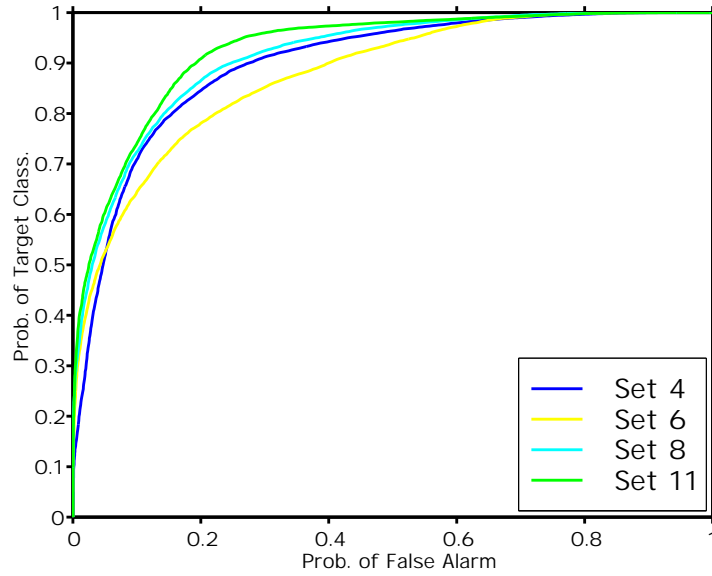
Figure 15: *The computed ROC curves for feature sets 4,6,8, and 11 using 246 swaths for training and 125 swaths for test (averaged over 301 realizations)*

Feature Set	FAR(80%)	FAR(90%)	FAR(95%)
1	0.329	0.542	0.663
2	0.223	0.346	0.538
3	0.239	0.370	0.525
4	0.155	0.277	0.430
5	0.342	0.534	0.675
6	0.222	0.401	0.533
7	0.186	0.290	0.420
8	0.142	0.247	0.379
9	0.153	0.280	0.423
10	0.146	0.257	0.389
11	0.127	0.191	0.267

Table 2: *False alarm rates for the various feature sets and for 80%,90% and 95% rates of target classification.*

Feature Set	error rate (KC)	NN1	NN3
1	0.234	0.271	0.273
2	0.179	0.237	0.237
3	0.189	0.237	0.241
4	0.157	0.224	0.220
5	0.241	0.328	0.306
6	0.197	0.217	0.207
7	0.160	0.203	0.216
8	0.148	0.170	0.191
9	0.159	0.197	0.203
10	0.150	0.183	0.198
11	0.121	0.205	0.205
12	0.138	0.137	0.176

Table 3: *Error rates obtained from kernel-based and nearest-neighbour classifiers*

shown in Table 4 with the error rates which were obtained using $N = 1$ and $N = 3$ nearest-neighbour classifiers during the same simulations. As can be seen, the error rate for the kernel-based classifiers are better for all feature sets except Set 12 which was constructed to optimized the nearest neighbour performance. However, it is interesting to note that for some of the feature sets, such as Set 8, the nearest neighbour performance is only slightly poorer than the kernel-based classifier. This was predicted from the curves of Fig. 12 which indicated that a small value of p also yielded a good error rate for feature set 8.

All these results were obtained by using 246 swaths for training and 125 for testing. In the results below, we use the same parameters, but use 370 swaths for training and one for testing. However, we sequentially test all possible combinations by using each swath in turn for testing with the remainder of the set for training. The parameters are not re-optimized for this bigger training set - we use the same values of p as for the smaller training set (246 swaths). Since the training set is now bigger, we would expect that for most of the classifiers that we should obtain at least as good results as before. This is always true for the 90% target classification rate; for the 80% classification rate the improvement was small or non-existent in some cases. For the 95% classification rate, the improvement was also smaller with the exception of Set 11 where the false alarm rate fell significantly from 26.7% to 19.7%. In Fig. 16 we show the ROC Curves for sets 4,6,8, and 11. As can be seen the resulting curves are not smooth and although Table 4 indicates that, for example, the classifier using set 8 is superiour to using set 4, it can be seen that within statistical uncertainty, that the 2 ROC curves for these feature sets appear to

Feature Set	FAR(80%)	FAR(90%)	FAR(95%)
1	0.287	0.508	0.656
2	0.221	0.295	0.443
3	0.246	0.320	0.467
4	0.139	0.213	0.393
5	0.262	0.500	0.631
6	0.213	0.369	0.508
7	0.197	0.238	0.393
8	0.139	0.205	0.369
9	0.139	0.238	0.418
10	0.131	0.213	0.344
11	0.115	0.172	0.197

Table 4: *False alarm rates for the various feature sets and for 80%,90% and 95% rates of target classification using all but one swath for training*

be very close.

For the results of Tables 2 and 3, we "tuned" the parameters, particularly p in the exponential kernel to obtain good results. We did investigate for Feature Set 8 (Features sets 1-4 combined with the template features) the classification performance obtained when first estimating the value of p from a validation set and then using that value to train for a test set. In this case we used 200 features for training, 100 for validation, and 71 for testing. We took a random partition of the swaths to obtain the training/validation set, estimated p^* from the validation set using 7 random partitions of the training/validation set, then retrained the classifier using the combined training/validation set with p^* and then tested on the testing set. This process was repeated 301 times and yielded an error rate of 0.145 (missed rate =0.063 and false alarm rate = 0.312) which is close to the value obtained when the tuned value was used in Table 2.

Instead of simply combining the features to make a large feature set, there are a variety of other techniques which can be used to combine the classification results from the feature sets. For example, we will consider using features sets 3,4, and 6. The default labels will be those from using the classifier with Set 4. However if both the discriminant values, w_3 and w_6 (i.e. the value from the function of Eq.(12)), from the classifier using sets 3 and 6, are such that both $w_3, w_6 < -t$ then we assign a label of -1 for the overall classification. Similarly if both values are great then $+t$, then a label of $= 1$ is assigned. In Fig. 17 the resulting error rate is shown as a function of t with a minimum error rate of 0.1534. The corresponding error rate for this simulation using just the set 4 classifier (using $\hat{b} =$ was 0.160, so we can see that by using
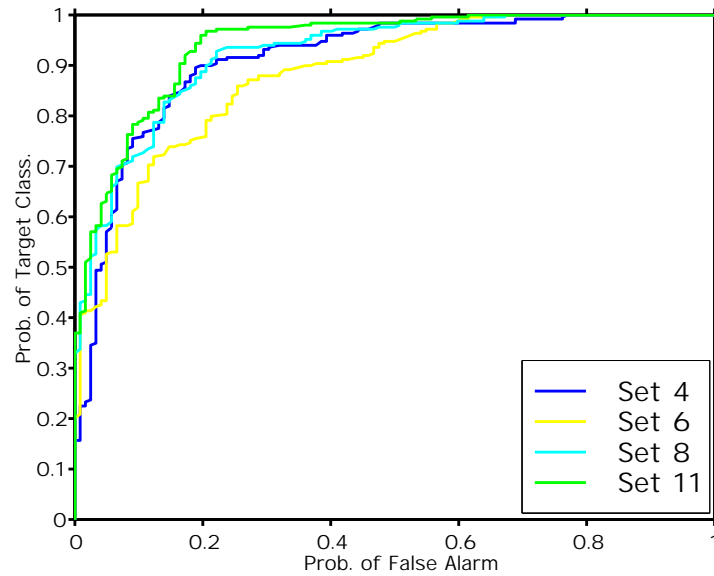
Figure 16: *The computed ROC curves for feature sets 4,6,8, and 11 using all but one swath for training*

Feature Set	error rate (KC)	NN1	NN3
1	0.213	0.248	0.267
2	0.148	0.216	0.213
3	0.170	0.213	0.208
4	0.132	0.210	0.199
5	0.205	0.323	0.291
6	0.175	0.218	0.197
7	0.143	0.186	0.191
8	0.121	0.142	0.178
9	0.143	0.189	0.202
10	0.129	0.159	0.167
11	0.089	0.189	0.183
12	0.108	0.094	0.140

Table 5: *Error rates obtained from kernel-based and nearest-neighbour classifiers using all but one swath for training*
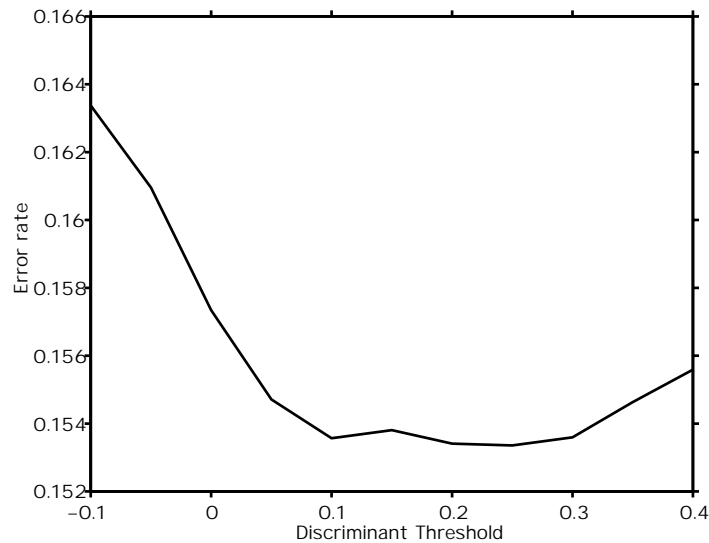
Figure 17: *The variation of classification error rate as the outputs from the classifier for Set 4 is combined with the outputs from the classifiers using Sets 3 and 6. The outputs from Set 3 and 6 are only used if both their discriminant values are below the negative discriminant threshold or above the discriminant threshold*

the outputs from the other 2 classifiers has improved the overall error rate by 4.0%. This does not seem to be better than combining the features from sets 1-4 and 6 as was done to construct Set 8, but it does provide another means for improving the performance of a classifier. There are, of course, many different ways the outputs from the classifiers using different features could be combined and we will study more of these in future work. For example, the concept of "bagging" and "boosting" [11] are methods for constructing a set of classifiers each of which concentrates on a portion of the overall classification process. In our case, it would also be reasonable to construct classifiers which were particularly good for a particular target type.

7   SUMMARY

The TTCP database of sidescan images is a challenging dataset for computer classification. There are several different target types with images from different locations. In this report we have considered the case of total automation. That is, the image initial segmentation and labelling of shadow and highlight regions is done without human guidance. Because of this, there are instances where any one particular set of segmentation parameters might fail because of the particular target/seabed conditions. In order to mitigate this problem somewhat, we defined 5 sets of segmentation parameters (and some differences in filtering). We gave examples of the shadow/highlight segmentations which resulted from various segmentation schemes for some of the swaths from the database. We also showed the distributions of height and length estimates which resulted for some of the target types of the database using the various filtering/segmentation methods. These distributions had a fairly significant standard deviation about the mean values. We also defined a set of model(template)-based features. We examined the classification performance which resulted when using these sets of features individually and in various combinations. It was found that using a kernel-based classifier yielded very good classification/false alarm rates. For some of the feature sets, the nearest-neighbour classifier also yielded very good classification results. We also showed that backwards feature selection could be used to determine optimized sets of features for both the exponential kernel-based and nearest-neighbour classifiers. For the case of 246 swaths for training and 125 swaths for testing, the optimized feature set for the kernel–based classifier yielded a false alarm rate of 19.1% for a 90% probability of target classification. Some of the feature sets were quite large so that a fairly large training set was required. However, these large sets did not seem problematic for the classifiers of this report. This is significant because it suggests that when the optimal set of filtering and segmentation parameters are not known apriori for a minehunting survey, one can simply combine the features obtained from a variety of filtering/segmentation methods without suffering a significant loss in performance. Another classifier strategy that was tried was to combine the outputs from the classifiers using the individual sets of features. This approach was successful in the sense that the resultant classification rates were, in general, superiour to using one of the single sets of features. However, this approach did not seem to be superiour to simply combining the individual sets of features into larger sets before the classification.

The database described in this report has proved to be a very valuable tool for providing a challenging testbed for classification methods. It is anticipated that we will continue to use this data for future studies and also acquire new

data to test against the classifiers of this report.

Appendix A: Details of features and filtering/segmentation parameters

A1. Non-template Features

These are features which are computed first by filtering the image, followed by segmentation, followed by region labelling and then perhaps some more filtering. Once the highlight and shadow regions have been defined, a variety of height, length and statistical measurements of the highlight and shadow regions are computed as features. These 38 features are now listed. Some of these features are also described in [3].

1. Feature 1 - Fourier Descriptor (1) - real part of shadow perimeter

2. Features 2-7- Fourier Descriptors (2-4) - real and imaginary parts.

3. Feature 8 Standard Deviation of shadow lengths (as function of along-track coordinate) normalized by mean value.

4. Feature 9 Number of pixels in designated shadow region

5. Feature 10 Area as defined by Feature 9 normalized by area of ellipse which fits shadow

6. Feature 11 Along-track length of shadow as defined by ellipse fit

7. Feature 12 Estimated height of target from shadow length using the maximum of the profile shadow lengths

8. Feature 13 Estimated height of target from shadow length using the maximum profile shadow length which is contained within shadow ellipse

9. Feature 14 Ratio of major axis/(minor axis +1) for fit ellipse (unity is included in denominator to prevent singularities)

10. Feature 15 Angle of ellipse that is fit to shadow

11. Feature 16 Length of target as estimated from leading edge of shadow profile - if along-track dimension is sufficiently large (> 7 along-track profile lengths are greater then 33% the maximum length) then across-track difference is included in computation.

12. Feature 17 A measure of the convexity of the shadow perimeter - the ratio of the perimeter of the shadow/perimeter of the convex hull

13. Feature 18 Along-track length of target (using the distance of those which are greater than 33% the maximum length

14. Feature 19 Lacunarity of pixel values within defined shadow area - the shadow area is redefined before the computation of lacunarity, so that the pixels between leading and trailing edges of shadow profile are also defined as shadow.

15. Feature 20 Area of designated highlight region - this is for the single connected region which is the highlight region [there are some simple criteria the highlight region must obey, such as the highlight is located before the shadow in across-track direction)

16. Feature 21 Mean along-track (divided by 2) length of highlight ellipse

17. Feature 22 Length of highlight using along and across-track coordinates from fit ellipse

18. Feature 23 Mean across-track (divided by 2) length of highlight ellipse

19. Feature 24 Aspect angle of fit ellipse to highlight

20. Feature 25 Ratio of number of pixels in highlight region to area of fit ellipse

21. Feature 26 Ratio of major axis to minor axis of ellipse

22. Feature 27 Lacunarity of highlight region

23. Feature 28 The next 6 features are the same as the previous 6 but now the highlight region is defined as all highlight clusters which lie within a specified region in front of the leading edge of the shadow. Area of this highlight region

24. Feature 29 Mean along-track (divided by 2) length of highlight ellipse

25. Feature 30 Length of highlight using along and across-track coordinates from fit ellipse

26. Feature 31 Mean across-track (divided by 2) length of highlight ellipse

27. Feature 32 Aspect angle of fit ellipse to highlight

28. Feature 33 Ratio of number of pixels in highlight region to area of fit ellipse

29. Feature 34 Ratio of major axis to minor axis of ellipse

30. Feature 35 Lacunarity of highlight region

31. Feature 36 The highlight region (using the definition of highlight for the previous 6 features) and shadow region are combined into one large region with pixels lying between the two regions also associated. The pixels of this region which are not in either the shadow or highlight

region are the counted (below, we refer to this region as the remainder region). The ratio of this number to the number of pixels in the combined region is the feature. It is hoped that this feature will be a measure of the transition zone between the highlight and shadow regions and also will measure any ringing or echo features within the shadow.

32. Feature 37 The number of pixels in this remainder region that are higher than 99.8% level mark of data.

33. Feature 38 the lacunarity of data values in this remainder region

Above, we have described the 38 features that are computed, all based upon the first determining shadow and highlight pixels and then performing a region labelling to define the shadow and highlight pixels to associate with the object. Thus the values of these features depend significantly upon the filtering and segmentation parameters that are used in the shadow and highlight determination. Below we describe the 5 sets of parameters which were used to compute the 5 sets of 38 features.

1. Method 1. - a 9-point median filter in the across-track direction is first applied to the data. For both shadow and highlight, the iterative segmentation method is used. For the definition of the shadow, the limits on the pixel values are quite restrictive. The starting value (requiring the pixel to be connected to at least 3 others) is the 3% level of the data and the upper value requiring connections to all other 8 is the 6% level. The highlight values start at the 99% level and go down to the 96% level. The determined highlight pixels are then median filtered with a 3-point across-track filter. The shadow pixels are median filtered with a 5-point median filter. The highlight and shadow masks are input into a region labelling algorithm. The two largest highlight and shadow regions are considered to determine the correct shadow/highlight pair based upon the sizes of each and the position of the highlight region relative to the shadow region.

   Another method of determining the highlight region is to consider all highlight regions whose centres are within a predefined distance of the leading edge of the shadow. This means several highlight regions can be considered as the target highlight. For this definition of highlight we used a more restrictive definition for the highlight segmentation and the non-iterative method was used. For all 5 methods, the highlight levels varied from 99.5% to 98%. However, it should be noted that the determined highlight regions will vary somewhat, as the leading edge of the shadow will differ for the different sets.

2. Method 2. - a 9-point median filter in the across-track direction is first applied to the data. For the shadow the non-iterative segmentation method is used and for the highlight the iterative method. The shadow values are now less restrictive than for Set 1 - varying from the 5% level for the lowest connectivity 3 to the 20% level for connectivity of at least 7. The highlight values start at the 99% level and go down to the 96% level which is the same as Set 1. The remainder of the operations is the same as Set 1.

3. Method 3. - a 5-point median filter in the across-track direction is first applied to the data. The iterative method is used for both the shadow and highlight. The shadow thresholds vary from a low value of 10% to the 30% level. The highlight values vary from a high of 99% to 96% level. The remainder of the operations are the same as the previous sets.

4. Method 4. - a 9-point median filter in the across-track direction is used. Another filter is applied in the along-track direction in which the data for each across-track index is normalized by the median of the data in the along-track direction for that index. This was done to help the segmentation in cases where the levels might be relatively high due to the local seabed and even the shadow values might have some speckle in them. This was often the case for the concrete cylinders in Esquimalt. The iterative method was used for both the shadows and highlight. The shadow thresholds are fairly "loose" with a starting level of 10% and a finishing level of 40%. The highlight levels vary from 99% to 95%. The remainder of the operations are the same as for previous Sets.

5. Method 5. - a 9-point median filter in the along-track direction is used. The shadow is determined by using a single hard threshold. This threshold is found by finding the threshold which minimizes the entropy of the image and we based our algorithm on the one developed for the SIPS segmentation [6]. The highlight used the iterative method and its levels varied from 99% to 97%.

A2 Template Features

In order to compute the template features, the swath image is first roughly segmented into highlight(positive values above the mean), background (zero), and shadow regions (negative values below the mean). This image is cross-correlated with simple target highlight/shadow templates representing the possible target types at various ranges, altitudes, and aspects(for non-symmetric

targets). In particular a ray-tracing code written in Fortran was used to compute the templates for the target classes: (1) Manta (2) vertical cylinder (taken to be 0.8 m high and 0.6 m diameter) (3) Mk 56 - cylinder 2.75 m long, 0.6 m diameter (4) MOG 5 cylinder 1.8 m long, 0.6 m diameter (5) Mk 62 cylinder 1.65 m long, 0.25 m diameter and (6) 1.7 m long, 42 cm diameter which represents the Mk 52 and 36 classes. These templates are computed for ranges (along the seabed) from 15 to 90 m in 5 m steps and for 3 different altitudes, 12,15, and 18 m. For the cylindrical objects 12 aspect angles are computed, from -90 degrees to 75 degrees in steps of 15 degrees. As discussed previously in the paper, for the cylindrical targets, the optimizing aspect is determined by maximizing the cross-correlation between the templates and the normalized image, with the cross-correlation normalized by the L_2 norm of the template. A set of 18 features are computed for each of the 6 template types, resulting in 108 features. The 18 basic features are listed below.

1. Feature 1 - Maximum cross-correlation of template with data, normalized by L_1 norm of template. In the case of cylindrical targets, the template is first matched with respect to aspect.

2. Feature 2 - Maximum cross-correlation of template with data, normalized by L_2 norm of template. In the case of cylindrical targets, the template is first matched with respect to aspect.

3. Feature 3 - Maximum cross-correlation of absolute value of template with absolute value data, normalized by L_1 norm of template. In the case of cylindrical targets, the maximizing aspect (determined on the basis of the standard cross-correlation) is used.

4. Feature 4 - Maximum cross-correlation of template with data, normalized by L_2 norm of template. In the case of cylindrical targets, the template is first matched with respect to aspect.

5. Feature 5 - The cross-correlation map between the template and the data is thresholded to 85% of the maximum value and the results region-labelled. The along-track length of the largest region is the feature.

6. Feature 6 - The cross-correlation map between the absolute value of template and absolute value of the data is thresholded to 85% of the maximum value and the results region-labelled. The along-track length of the largest region is the feature.

7. Feature 7 - The number of shadow pixels in the image which fall in the shadow portion of the template normalized by the number of shadow pixels in the template.

8. Feature 8 - The number of highlight pixels in the image which fall in the highlight portion of the template normalized by the number of highlight pixels in the template.

9. Feature 9 - The number of shadow pixels in the image which fall in the highlight portion of the template normalized by the number of shadow pixels in the template.

10. Feature 10 - The number of highlight pixels in the image which fall in the shadow portion of the template normalized by the number of highlight pixels in the template.

11. Feature 11 - The number of non-zero pixels of the absolute value of image which overlap the absolute value of the template normalized by the number of non-zero pixels in the absolute value of the template.

12. Feature 12 - The normalized image data is smoothed with a 25-point two-dimensional Gaussian filter (i.e., 5x5 pixels) and then differenced in the along-track direction. The same operations are performed upon the template and the cross-correlation is computed.

13. Feature 13 - The normalized image data is smoothed with a 25-point two-dimensional Gaussian filter (i.e., 5x5 pixels) and then differenced in the across-track direction. The same operations are performed upon the template and the cross-correlation is computed.

14. Feature 14 - The normalized image data (absolute value) is smoothed with a 25-point two-dimensional Gaussian filter (i.e., 5x5 pixels) and then differenced in the along-track direction. The same operations are performed upon the absolute value of template and the cross-correlation is computed.

15. Feature 15 - The normalized image data (absolute value) is smoothed with a 25-point two-dimensional Gaussian filter (i.e., 5x5 pixels) and then differenced in the across-track direction. The same operations are performed upon the absolute value of the template and the cross-correlation is computed.

References

1. G.J. Dobeck, J.C. Hyland and L. Smedley, "Automated detection/classification of seamines in sonar imagery" Proc. SPIE-Int. Soc. Optics, vol. 3079, pp.90-110, 1997.

2. I. Quidu, Ph. Malkasse, G. Burel, and P. Vilbe, "Mine classification based on raw sonar data: An approach combining Fourier descriptors, statistical models and genetic algorithms" OCEANS MTS/IEEE Conf. And Exhibition, Vol. 1, pp.285-290, 2000.

3. V.L. Myers, "Decision trees for computer-aided detection and classification (CAD/CAC) of mine in sidescan sonar imagery" DRDC Atlantic TM 2002-144, December 2002.

4. P.C Connor, "Identifying features that distinguish between mine-like objects in sidescan sonar imagery", Thesis, Master of Science in Engineering, University of New Brunswick, 2002.

5. J.A. Fawcett, "Computer-aided detection and classification of minelike objects using template-based features", Oceans 2003, San Diego, September 2003.

6. Interim Report - Development and Implementation of classification tools with sidescan sonar image database, DRDC Atlantic Contract, MacDonald Dettwiler and Associates, April, 2004.

7. M.A. Trevorrow, A. Crawford, J. Fawcett, R. Kessel, T. Miller, V. Myers and M. Rowsome, "Synopsis of survey data collected during Q-260 MAPLE 2001 sea-trials with CFAV Quest and NRV Alliance", DREA TM 2001-172, November, 2001.

8. V.L. Myers, *Image segmentation using iteration and fuzzy logic* in CAD/CAC 2001 Conference Proceedings, 2001.

9. N. Cristianini and J. Shawe-Taylor, *An introduction to support vector machines and other kernel-based learning methods*, Cambridge University Press, 2000.

10. T.an Gestel, J. Suykens, G. Lanckriet, A. Lambrechts, B. De Moor and J. Vandewalle, "Baysian framework for least squares support vector machine classifiers, Gaussian Processes, and Kernel Fisher discriminant analysis", Maunscript no. 2281, Dept. of Electrical Engineering ESAT-SISTA, Katholieke Universiteit Leuven, 2001.

11. R. Duda, P. Hart, and D. Stork, *Pattern Classification, Second Edition*, John Wiley and Sons, New York, 2001.

12. IMSL Fortran Subroutines for Mathematical Applications, Visual Numerics Inc., 1997

Distribution List

Internal Distribution

David Hopkin,
Mine and Torpedo Defence Group,
9 Grove St.,
Dartmouth, Nova Scotia, B2Y 3Z7

Anna Crawford,
Mine and Torpedo Defence Group,
9 Grove St.,
Dartmouth, Nova Scotia, B2Y 3Z7

Juri Sildam,
Mine and Torpedo Defence Group,
9 Grove St.,
Dartmouth,Nova Scotia, B2Y 3Z7

John Fawcett,
Mine and Torpedo Defence Group,
9 Grove St.,
Dartmouth,Nova Scotia, B2Y 3Z7

Ron Kessel,
Maritime Information Knowledge Management Group,
9 Grove St.,
Dartmouth,Nova Scotia, B2Y 3Z7 Library (5)

External Distribution

NDHQ/DRDKIM

Vincent L. Myers,
NATO Undersea Research Centre,
Viale San Bartolomeo 400,
19138 Ls Spezia (SP),
Italy

Dr. Gerald Dobeck,
NSWC Coastal Systems Station
Dahlgren Division, Code R24

6703 Hwy 98
Panama City, FL 32407-7001, U.S.A.

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's report, or tasking agency, are entered in section 8.) Defence R&D Canada – Atlantic, PO Box 1012, Dartmouth, NS B2Y 3Z7		2. SECURITY CLASSIFICATION (overall security classification of the document including special warning terms if applicable). UNCLASSIFIED

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S,C,R or U) in parentheses after the title). Computer-aided classification for a database of images of minelike objects

4. AUTHORS (Last name, first name, middle initial. If military, show rank, e.g. Doe, Maj. John E.) John A. Fawcett

5. DATE OF PUBLICATION (month and year of publication of document) March 2005	6a. NO. OF PAGES (total containing information Include Annexes, Appendices, etc). 38 (approx.)	6b. NO. OF REFS (total cited in document) 12

7. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered). Technical Memorandum

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include address). Defence R&D Canada – Atlantic PO Box 1012 Dartmouth, NS, Canada B2Y 3Z7

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant). 11CL	9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written).

10a ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.) DRDC Atlantic TM 2004-272	10b OTHER DOCUMENT NOs. (Any other numbers which may be assigned this document either by the originator or by the sponsor.)

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification) (X) Unlimited distribution (   ) Defence departments and defence contractors; further distribution only as approved (   ) Defence departments and Canadian defence contractors; further distribution only as approved (   ) Government departments and agencies; further distribution only as approved (   ) Defence departments; further distribution only as approved (   ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected).

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual).

In this report we describe some computer classification experiments with a database of sidescan sonar images. This database consists of 383 swaths of sidescan sonar data extracted by the authors from sea trial data collected over the last few years by DRDC Atlantic. The effects of the filtering and image segmentation processes on the resultant classification rates are considered. A number of kernel-based and nearest-neighbour classification schemes are examined. It is found that despite the complexities of the database considered in this report that high classification/low false alarm rates can be achieved.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title).

sidescan sonar, classification

This page intentionally left blank.

Defence R&D Canada

Canada's leader in defence
and National Security
Science and Technology

R & D pour la défense Canada

Chef de file au Canada en matière
de science et de technologie pour
la défense et la sécurité nationale

DEFENCE R&D **DÉFENSE**

www.drdc-rddc.gc.ca